第 38 卷第 11 期
2018 年 6 月

生 态 学 报
ACTA ECOLOGICA SINICA

Vol.38,No.11
Jun.,2018

DOI: 10.5846/stxb201705080841
董仁才,王韬,张永霖,张雪琦,李欢欢.我国城市可持续发展能力评估指标的元数据分析与管理.生态学报,2018,38(11):3775-3783.
Dong R C, Wang T, Zhang Y L, Zhang X Q, Li H H.Metadata analysis and management for an urban sustainable development capability evaluation index in China.Acta Ecologica Sinica,2018,38(11):3775-3783.

我国城市可持续发展能力评估指标的元数据分析与管理

董仁才[1,*],王 韬[1,2],张永霖[1,2],张雪琦[1,2],李欢欢[1,2]
1 中国科学院生态环境研究中心城市与区域生态国家重点实验室,北京 100085
2 中国科学院大学,北京 100049

摘要:在我国大力推动城市可持续发展,推进国家可持续发展实验区建设的同时,采用何种评估方法和数据开展城市可持续发展能力评估是需要重点解决的问题。近年来兴起的元数据理论与技术在解决评估数据质量控制方面被视为是一种行之有效的方法。针对我国现阶段使用的一些城市可持续发展能力评估指标体系的特点,通过深入剖析每一个指标数据的来源、获取手段、适用方法等特征,提出从软件工程学思路研发城市可持续发展能力评估元数据管理系统的具体方法,帮助可持续发展实验区高效获取和管理评估所需数据信息;以"十二五"科技支撑计划项目"城市可持续发展能力评估及信息管理关键技术研究与示范"中所建立的元数据规范,对其所包含的"数据发布日期"、"数据发布形式"、"空间范围"、"时间范围(起始时间、结束时间)"、"统计频率"、"数据安全限制分级"、"数据志说明"、"在线资源链接地址"和"数据统计单位信息(单位名称、联络人、联系电话、单位地址、邮件地址)"共 14 项为评估数据的关键元数据项,以此追踪对标的评估数据。并通过量化数据质量评分法针对数据质量在运用元数据追踪法前后的评价结果对比发现,被评估指标的数据质量在获得元数据支持时,其数据可靠性、可比性和可持续性方面的评价分值都获得了十分显著的改善。研究认为采用元数据理论在控制和保障城市可持续发展能力评估数据质量方面具有优势作用,开发有针对性的城市可持续发展能力评估元数据管理系统能够有效提高评估数据的综合评价结果。

关键词:城市可持续发展能力评估;元数据管理系统;数据质量评价;可持续发展实验区

Metadata analysis and management for an urban sustainable development capability evaluation index in China

DONG Rencai[1,*], WANG Tao[1,2], ZHANG Yonglin[1,2], ZHANG Xueqi[1,2], LI Huanhuan[1,2]
1 *State Key Laboratory of Urban and Regional Ecology*, *Research Center for Eco-Environmental Science*, *Chinese Academy of Sciences*, *Beijing* 100085, *China*
2 *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*

Abstract: The scientific assessment of urban sustainable development capability has important guiding significance for the urban green development and smart growth. Although China is promoting urban sustainable development and developing sustainable experimental areas, it is vital to identify scientific methods and obtain data to analyze this development. In recent years, metadata has been regarded as an effective method to improve data quality. This study focused on the characteristics of an urban sustainable development analysis index system which is based on a metadata management system using software engineering by analyzing the data source, collection method, and application method etc. The aim of this study was to effectively and efficiently collect useful data from the sustainable development experimental area. According to

基金项目:国家重点研发计划(2016YFC0503605)
收稿日期:2017-05-08; **网络出版日期**:2018-03-02

* 通讯作者 Corresponding author.E-mail: dongrencai@rcees.ac.cn

the metadata standard which was developed by "research and demonstration on urban sustainable development capacity assessment and information management key technologies" in 12th Five-Year science and technology support program, we used fourteen key evaluation metadata items in this standard as benchmarks, including data release date, data release form, spatial range, time range (start time and end time), statistical frequency, data security limit classification, data description, Online resource link address and information of data statistics institution (institution name, contacts, contact number, address and e-mail address). Comparing the results of data quality evaluation before and after using a metadata tracking method through the quantitative data quality scoring method, we found that the data are more reliable, comparable and constant when using metadata. This research shows several advantages in controlling and protecting urban sustainable development analysis data quality when applying the metadata theory. Further, the development of a specific urban sustainable development analysis metadata management system could improve the comprehensive evaluation result of the data.

Key Words: urban sustainable development capacity; metadata management system; data quality evaluation; sustainable development experimental zone

城市可持续发展能力评估指标体系作为一种有效的评估工具,近年来已经从理论研究走向了实际应用[1-4]。虽然目前还没能形成一套公认的或者被普遍接受的评价方法及指标体系[5],但就目前各类评价指标体系的实际应用而言,对评估体系所需的统计数据进行有效管理,提高数据质量是保障评估结果精准性、可取性和实用性的重要方式。统计数据质量概念是质量的一般涵义与统计工作及其结果具体相结合而产生的科学范畴[6]。这意味着数据质量不仅包括数据本身,还包括了数据的生产与实现过程以及完成这一过程中的工作质量。因此,对统计数据质量的控制不应仅限于对数据本身的准确性要求,还应囊括可靠性、可比性、可得性、适用性、有效性和方法专业性等不断扩大的统计数据质量内涵[7]。而评估结果是否能准确、及时的反应评估对象的可持续发展能力,并给予决策者具备参考价值的适用性信息,先决条件就在于提高数据质量,尤其是完善包含多种数据需求的数据质量。然而由于城市可持续发展能力基于资源、社会、经济、环境等各子系统上,评估数据源于多个不同的数据生产部门,复杂的数据来源使得在满足数据多维性要求、提高评估数据质量方面变得更加困难[8-9]。

元数据是关于数据的信息,是具有描述、解释、定位信息资源功能的结构化信息,是说明数据内容、质量、状况及其他有关特征的描述信息[10]。对元数据的相关研究始于数字图书管理和网络搜索引擎开发领域[11],其目的是为了解决网络信息飞速发展下的数据量骤增现象[12]。元数据作为数据背后的记录者,在多元化呈现数据本身内涵方面有极为优秀的表现。当前评估数据质量不高的一个重要因素就是没有重视对数据背后关键信息的有效记录和管理,而元数据正是帮助"发现、识别、确认、记录"数据背后的关键信息,从而完成"管理,检验、使用和保障"评估数据的基本功能,诸如记录数据来源渠道、透明化数据收集过程、规范化数据质量责任制、保障数据真实性和准确性。

在城市可持续发展能力评估数据质量亟待提高、促使评估成果能更加全面精准的反映评估对象可持续发展能力的迫切要求下,本文通过对当前评估数据的来源分析,基于元数据理论在保障数据质量方面的优势作用,通过追踪城市可持续发展评估指标数据的关键元数据,运用数据质量评分法对城市可持续发展能力评估指标的数据质量在获取相关元数据前后的变化进行评分,并以此研发和使用城市可持续发展能力评估元数据管理系统,帮助可持续发展实验区高效获取和管理评估所需数据信息,为解决城市可持续发展能力评估中的数据问题提供方法,保证评估结果科学有效。

1 研究方法

1.1 评估数据的来源分析

城市可持续发展能力评估的主要研究热点多集中于评估指标体系和评估方法的设计与创新方面,但评估

数据质量的高低也是影响可持续发展能力评估结果科学有效的关键因素。城市可持续发展能力评估一直具有很大的特殊性和复杂性,城市间不同的自有发展模式、城市规模的大小[13]、以及所用评估体系的差异性,都使得评估工作的数据来源更为复杂。

根据国家“十二五”科技支撑计划项目“城市可持续发展能力评估及信息管理关键技术研究与示范”中的《城市可持续发展能力评估指标体系》(以下简称评估指标体系)的数据来源分析表明,其所含的 49 个评估指标涉及 71 处数据来源,涵盖了 54 个相关部门的统计数据资料(图 1)。其中,统计年鉴资料占比 60%,统计公报资料占比 34%,其他来源资料占比 6%。城市可持续发展能力评估要素涉及面广、涵盖内容丰富,而如此之全面的统计数据只有相关统计年鉴及统计公报才能满足。其次,我国统计法规定只有政府认可及发布的相关经济社会数据才具备公信力,也即基于此的城市可持续发展能力评估结果才具有认可度。因此,统计年鉴及相关统计公报作为最常见的政府数据公布方式,成为评估数据的主要获取来源。

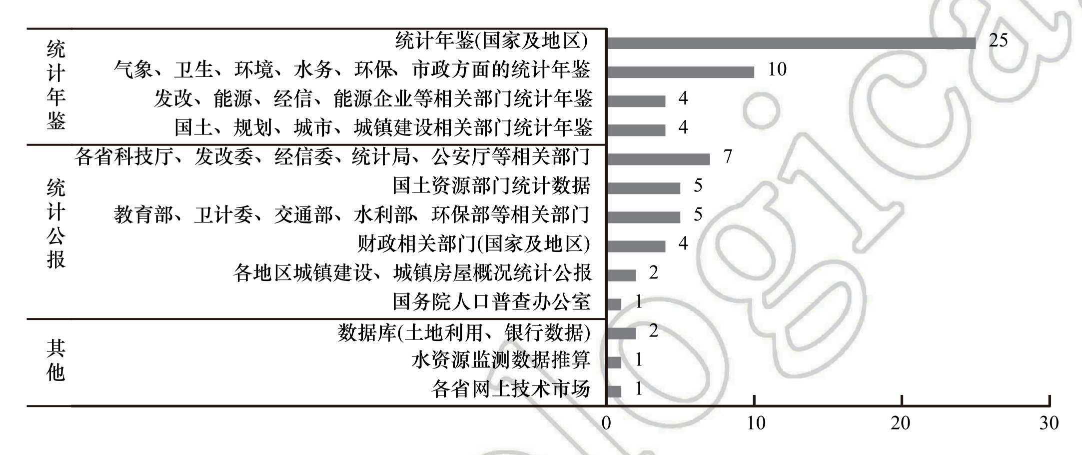


图 1　城市可持续发展能力评估指标体系数据来源统计图

Fig.1　The data source statistical chart of urban sustainable development capability evaluation index

1.2　数据质量评价方法

1980 年,联合国统计局(United Nations Statistical Office)出版的《统计组织手册》对各国的官方统计资料提出了八项要求:“统计资料应考虑使用者的需求;统计服务应针对多种使用者;统计资料应建成有机联系体系;统计资料应及时收集发布;统计资料应保持历史连续性;应充分保障被调查者的利益或秘密;统计机构应保持客观公正;统计机构必须有胜任统计工作的行政能力”。至此,统计数据的质量概念开始由狭义的仅包含数据准确性,向广义的包含数据准确性、及时性、可比性、可得性、适用性、有效性和方法专业性等多维性的内涵过度[14]。

因此,为了直观展示城市可持续发展能力评估的数据质量,本文建立针对评估指标体系所需数据的质量评价方法。该评价方法以数据质量多维性要求为指导,设计按照数据质量的可靠性、可比性和可持续性进行分类评价,并综合得出指标体系所需数据质量的评价结果[15-16]。同时,鉴于当前评估工作数据来源普遍为官方统计数据,而统计年鉴作为政府机构的数据主要发布形式,因而该评价方法主要以统计年鉴数据的质量为评价目标。其中,可靠性信息主要包括数据来源的统计年鉴类别和数据公开情况;数据可比性信息主要包括数据的计算方法、适用范围和统计频率;数据可持续性信息主要包括数据的缺失频率、数据时间范围和历史数据的收集方式。具体的量化评价体系如下所示。

(1)数据质量可靠性评价。从数据来源的权威性和数据的公开情况来评价指标体系中数据质量的可靠性。其中,数据来源权威性可以分为国家级、地市级和部门级。国家级定义为数据获取渠道由国家统计局及相关部委的统计部门发布,地市级定义为由地方政府或地方统计局发布(省,市),部门级定义为地方单位统计部门出具的统计年报或无法获取相应统计数据。数据公开情况可以分为公开、部分公开和不公开,其中公开定义为可直接查询到的数据,部分公开定义为需要联系相应机构获取的数据,不公开定义为无法获取的数

据。其可靠性相应得分分别为 10、9 和 8,具体见表 1。

表 1 数据质量可靠性评分

Table 1 Scores of data quality reliability

数据来源权威性 Data source of authority	得分 Score	数据公开情况 Data disclosure situation	得分 Score
国家级 National level	10	公开	10
地市级 Prefectural level	9	部分公开	9
部门级 Department level	8	不公开	8

(2)数据质量可比性评价。从数据的适用范围和生产方式来评价上述指标体系中数据质量的可比性。其中,实验区类型可以分为所有地区、已核实地区和未核实地区,其中所有地区定义为数据适用空间范围为全国各个地区,已核实地区定义为数据适用空间范围所属省或市县,未核实地区定义为数据适用空间范围不明。生产方式可以分为直接获取、间接计算和理论推算,其中直接获取定义为数据已有官方认证的对应结果可直接使用,间接计算定义为对应结果需要自行收集官方认证的相关数据计算获取,理论推算定义为对应结果由无官方认证的相关数据或文献类数据计算获取。其可比性相应得分分别为 10、9 和 8,具体见表 2。

表 2 数据质量可比性评分

Table 2 Scores of data quality comparison

适用范围 Application range	得分 Score	生产方式 Production mode	得分 Score
所有地区 All regions	10	直接获取	10
已核实地区 Verified area	9	间接计算	9
未核实地区 Unverified area	8	理论推算	8

(3)数据质量可持续性评价。从数据的时间范围和收集方式来评价上述指标体系中数据质量的可持续性。其中,时间范围可以分为 0—5 年、6—10 年和 11—15 年,其中 0—5 年定义为近 5 年内有连续数据可查或无法查询,6—10 年定义为近 6 到 10 年有连续数据可查,11—15 年定义为近 11 到 15 年有连续数据可查。数据的收集方式可以分为全部来自发布数据、部分来自发布数据和非发布数据,其中全部来自发布数据定义为数据采集直接来自官方公开发布的数据,部分来自发布数据定义为数据采集直接来自官方已公开发布的数据和未公开发布的相关政府机构数据,非发布数据定义为数据采集来自未公开发布的相关政府机构数据。其得分相应分别为 10、9 和 8,具体见表 3。

表 3 数据质量可持续性评分

Table 3 Scores of data quality sustainability

时间范围 Time range/a	得分 Score	收集方式 Collection mode	得分 Score
11—15	10	全部来自发布数据	10
6—10	9	部分来自发布数据	9
0—5	8	非发布数据	8

本文运用典型抽样法抽取评估指标体系中的 21 个指标进行数据质量的量化评分,但该评价方法仍然适用于整套评估指标体系的数据质量评价。评价选取中国统计局官方网站数据(http://www.stats.gov.cn/)、《中国城市统计年鉴》数据(1996—2015)、《中国环境统计年鉴》数据(1998—2015)定义为国家级;选取案例城市绍兴市统计信息网数据、《绍兴市统计年鉴》数据(2000—2015)定义为地方级;选取《绍兴市科技信息研究院统计报表》(2010—2015)定义为部门级。

1.3 元数据追踪法

元数据是关于数据的信息,是具有描述、解释、定位信息资源功能的结构化信息,是说明数据内容、质量、状况及其他有关特征的描述信息[10]。当前评估数据质量不高的一个重要因素就是没有重视对数据背后关键信息的有效记录和管理,而元数据正是帮助“发现、识别、确认、记录”数据背后关键信息,从而完成“管理,检验、使用和保障”评估数据的基本功能,可见元数据理论正是解决提升数据质量可靠性、可比性和可持续性的有效方法。

本文参照“十二五”科技支撑计划项目“城市可持续发展能力评估及信息管理关键技术研究与示范”中所建立的元数据标准,以其所包含的“数据发布日期”、“数据发布形式”、“空间范围”、“时间范围(起始时间、结束时间)”、“统计频率”、“数据安全限制分级”、“数据志说明”、“在线资源链接地址”和“数据统计单位信息(单位名称、联络人、联系电话、单位地址、邮件地址)”共 14 个元数据项为评估数据的关键元数据项,以此追踪对标的评估数据。同时运用数据质量评价方法对评估指标体系的数据质量在使用元数据追踪前后的变化进行评分,以期检验元数据在保障数据质量方面是否具有优势作用。

2 研究结果与分析

2.1 评估数据质量的现状分析

通过对大量可持续发展实验区调研分析表明,当前国内城市可持续发展能力评估中的数据质量存在一定问题,直接或间接的对评估结果产生影响。其具体表现在:

(1)数据可靠性不足。城市可持续发展能力评估数据的产生部门众多,且所需数据不全为公开发布的数据,相当一部分非公开数据需要评估主管部门联系相关数据产生部门获取,兼之获取后没有系统记录这些数据的获取信息,数据问题又难以及时溯回和纠正,阻碍了评估工作进行,也同样为来年的评估数据采集加重负担。

(2)数据的可比性不足。我国地域广阔,城市类型多样。不同气候类型区,不同功能特点的城市对同一评估指标的统计口径理解上也有差异,同时由于评估工作缺乏对数据产生方式的信息记录,使缺乏可比性的数据应用于评估中,因而数据可比性不足已经成为推动不同城市间进行可比性研究和全面落实可持续发展战略的阻碍。

(3)数据的可持续性不足。可持续发展能力评估工作的长期性决定了评估数据的获取必须要有持续性。然而,评估过程中的历史数据难以追溯引发的时间尺度评价断层,严重影响了评估结果的科学性与准确性。同时,面对覆盖全国的不同空间、不同地域、不同行政级别的可持发展实验区和多源化的数据生产部门,如果没有对数据进行持久且完备的记录,亦难以保障数据的真实性和准确性。

以上这些,都对评估结果的适用性造成直接或间接的影响。而目前使用的数据获取及管理机制中,并没有针对数据质量内涵进行扩展补充,这种对数据质量内涵进一步完善的缺失是导致评估中产生数据问题的根本原因。

2.2 元数据追踪后数据质量对比分析

根据前述数据质量评分标准,对所选指标的评估数据在追踪其关键元数据项前后的数据质量进行打分,其结果如表 4。

从表 4 可以看出,各项指标的数据质量评分在追踪其对应的元数据项后,都获得了十分显著的改善,并且数据质量在可靠性、可比性和可持续性方面的评价分值都获得了相应的提高。而数据质量新评分值的整体提高,得益于元数据完善了评估数据的来源信息,基础信息,统计方式及口径信息。在来源信息方面:数据发布时间及发布形式信息的收录,为该项指标的数据收集工作指明方向,降低了原指标数据采集过程中的不确定性,提高了数据质量的可持续性。数据生产单位信息被收录,包括单位名称、负责人、电话地址等,做到数据责任到人,便于数据出现问题后的及时回溯与纠错,提高了数据质量可靠性;在基础信息方面:数据志说明进一

步对评估数据做补充解释,如计算公式、计算所需的相关数据获取源及固定元数据项不足以描述的相关信息等,提高指标数据质量的可比性;在统计方式及口径信息方面:包括统计频率、统计时间及空间范围信息的收录,保障数据在不同地区使用时的可比性。由此可见,追踪评估数据的关键元数据项对数据质量的提升是全面性的,加强对评估数据相关元数据的记录和管理在长效保障评估数据质量方面有着不可或缺的优势。同时,将评估数据的元数据追踪采集制度化和长效化,使数据记录保持在长期稳定的收录状态就是可以预期的,数据记录将达到甚至打破数据质量可持续性评分的上限,从而满足城市可持续能力评估工作的长期性需求,提高数据质量的可持续性。

表 4 基于元数据追踪的数据质量评分结果对比

Table 4 Comparison for score results of data quality according to metadata management system

编号 Number	指标 Index	数据可靠性平均分(原/新) Average of data quality reliability (old/new)	数据可比性平均分(原/新) Average of data quality comparison (old/new)	数据可持续性平均分(原/新) Average of data quality sustainability (old/new)	总得分平均分(原/新) Average of total scores (old/new)
1	生态系统服务价值总量	8.50/9.50	8.00/8.50	8.00/9.50	8.17/9.17
2	突发环境污染事件发生次数	8.50/9.50	9.00/9.00	8.50/10.0	8.67/9.50
3	温室气体排放强度	8.50/9.50	8.50/9.00	8.00/10.0	8.33/9.50
4	废水排放强度	9.00/9.50	9.00/9.50	9.00/10.0	9.00/9.67
5	废弃物排放强度	9.00/9.50	9.00/9.50	9.00/10.0	9.00/9.67
6	主城区人均公共绿地面积	8.50/9.50	8.50/9.00	9.00/10.0	8.67/9.50
7	环保投资占财政支出比重	8.50/9.50	8.50/9.00	8.50/10.0	8.50/9.50
8	污水排放达标率	8.50/9.50	8.50/9.00	9.00/10.0	8.67/9.50
9	工业固废综合利用率	9.00/9.50	9.00/9.50	9.50/10.0	9.17/9.67
10	年空气质量达标日	8.50/9.50	9.00/9.50	9.00/9.50	8.83/9.50
11	万元 GDP 温室气体排放当量	8.50/9.50	8.50/9.50	9.00/10.0	8.67/9.67
12	万元 GDP 水耗	8.5/09.50	9.00/9.50	8.50/10.0	8.67/9.67
13	城市紧凑度	8.50/9.50	8.00/8.50	8.50/9.50	8.33/9.17
14	城镇单位建设用地产值	8.50/9.50	8.50/9.50	8.50/10.0	8.50/9.67
15	人均耕地面积	10.0/10.0	8.50/9.50	9.50/10.0	9.33/9.83
16	人均建设用地面积	8.50/9.50	9.00/9.50	9.00/10.0	8.83/9.67
17	人均水资源可利用量	8.50/9.50	8.50/9.50	9.50/10.0	8.83/9.67
18	能源对外依存度	9.00/9.50	8.50/9.00	8.50/10.0	8.67/9.50
19	非化石能源消耗比重	8.50/9.50	8.50/9.00	9.00/10.0	8.67/9.50
20	城镇燃气普及率	8.50/9.50	9.00/9.00	8.50/10.0	8.67/9.50
21	城镇供水保障能力	8.50/9.50	8.50/9.00	8.50/10.0	8.50/9.50

最后,为了对追踪评估数据关键元数据项是否能提高数据质量进行科学有效的评价,下表对数据质量评分结果进行了相关的统计描述和 T 检验(表 5),分别从数据的可靠性、可比性、可获得性以及总评这 4 个方面来说明元数据管理系统的效用。

(1)数据的可靠性上,追踪关键元数据让数据质量的评分得到了整体上升。这得益于在实验区实行的元数据采集一套表模式,使数据统计部门在数据生产及发布过程中做到互通有无、步调一致,改变过去数出多门、数据打架的纷乱局面,使数据来源权威性评分指标从部门级上升到了地方级及以上;同时,广泛公开的元数据项追踪及管理提升了数据的透明性,其数据来源责任到人的管理方式,将杜绝数据生产部门再行无中生有、改头换面的统计造假,也将督促生产部门快速发布数据,提高数据的使用效率,使得元数据一套表的获取方式成为实验区评估数据的有效管理保障措施。因此,数据可靠性评价获得了 9.5 以上的优于原始数据的评分,获得元数据支撑后的评估数据在可靠性这一指标上的得分更高,较原始数据提升了 9.8%。同时,基于 T

检验的结果显示,数据质量可靠性的提高是极显著的(P<0.01),对数据质量可靠性的提高是整体上升的,因而追踪评估数据的关键元数据使得数据在可靠性上会大幅优于其未获元数据支持的评估数据。

(2)数据的可比性上,鉴于追踪的元数据标准在数据统计口径上的一致性表述,使绍兴市范围内生产的统计数据在统计范围、统计方式等方面获得统一性规范,改变了数据生产部门各自为政的统计现状,保证了数据在系统使用范围内的可比性。甚至可以预计的是,在元数据追踪一套表模式向全国 180 多个国家可持续发展实验区全覆盖时,将能极大地提高各实验区可持续发展能力评估结果的比对性,推动我国在全面了解各地区可持续发展能力优缺性的基础上做到更为科学合理的发展布局。同时,经 T 检验的结果可知,获得元数据支持的数据在可比性上也有极显著(P<0.01)的提升,且评价结果显示新评分均值比原始评分均值提高了 0.55,表示数据质量的可比性同样获得了明显的提升。

(3)数据的可获得性上,由于元数据追踪在保持评估数据长效获取方面的可预计性,因此在保持元数据的完备记录期内能提高评估数据质量的可持续性。T 检验结果方面,新评分均值比原始评分的均值提高了 1.14,增幅达 13%,在所有指标中变化最为显著。同时,新评分结果总体标准差更小说明数据稳定性更强,相较于旧的评分结果而言,数据的可获取性得到了更高水平的提升。

(4)在数据质量的总评分上,新评价的综合得分为 9.55,高于原评价 0.85 分,总体增幅约为 9.8%。评价结果显示元数据对数据质量的提升作用优势明显(P<0.01),且对比新旧评分结果的标准差可以看出,获得元数据支持的评估数据,其数据质量提升表现的更为稳定。

表 5　数据质量评分统计结果

Table 5　Statistical results of data quality

评价项目 Evaluation items	新评分 New scores		原评分 Old scores		差值 Difference value	T 值 T-value
	均值 Average	标准差 Sample variance	均值 Average	标准差 Sample variance		
数据可靠性 Data reliability	9.52	0.11	8.67	0.36	0.85	14.01
数据可比性 Data comparability	9.19	0.33	8.64	0.31	0.55	9.31
数据可获得性 Data availability	9.93	0.17	8.79	0.42	1.14	12.39
总评分 Total score	9.55	0.16	8.70	0.27	0.85	21.40

3　城市可持续发展能力评估元数据管理系统

综上所述,完善评估数据的元数据在提高评估数据质量方面的优势明显。因此,本研究结合“十二五”科技支撑计划项目“城市可持续发展能力评估及信息管理关键技术研究与示范”研究工作,开发基于评估数据元数据项的管理系统,高效便捷的记录和管理评估数据的元数据信息,保障城市可持续发展能力评估工作中的数据质量。该系统是对描述数据的数据进行集中存储并管理的数据库管理系统。城市可持续发展元数据管理系统的系统物理结构为三层架构(Three-tier architecture),具体分为用户界面层(User interface layer)、业务逻辑层(Business logic layer)和数据访问层(Data access layer)3 个部分。系统严格按照软件工程的方法和流程进行开发,采用 B/S(Browner/Server)[17-18] 架构设计了一个面向用户的元数据交互平台,旨在为数据提供方、数据审核方以及最高级管理人员提供一套元数据管理的解决方案,元数据管理系统的总体组织结构如图 2。

元数据管理系统采用了用户分级制,将用户划分为数据提供方、数据审核方以及最终验收方三类。三者拥有的系统管理和应用权限依次递增,即最终验收方拥有最高级的数据管理和检索权限,相反,数据提供方只可以提交和管理地方所属城市的元数据信息。权限分级一方面符合政府行政管理的规范化流程,另一方面便于数据审核和管理的权责到人,从而达到严格控制元数据质量的目的[19]。另外,用户分级还利于该系统的开发,功能的模块化和等级化有利于数据的安全保护和元数据库管理[20]。

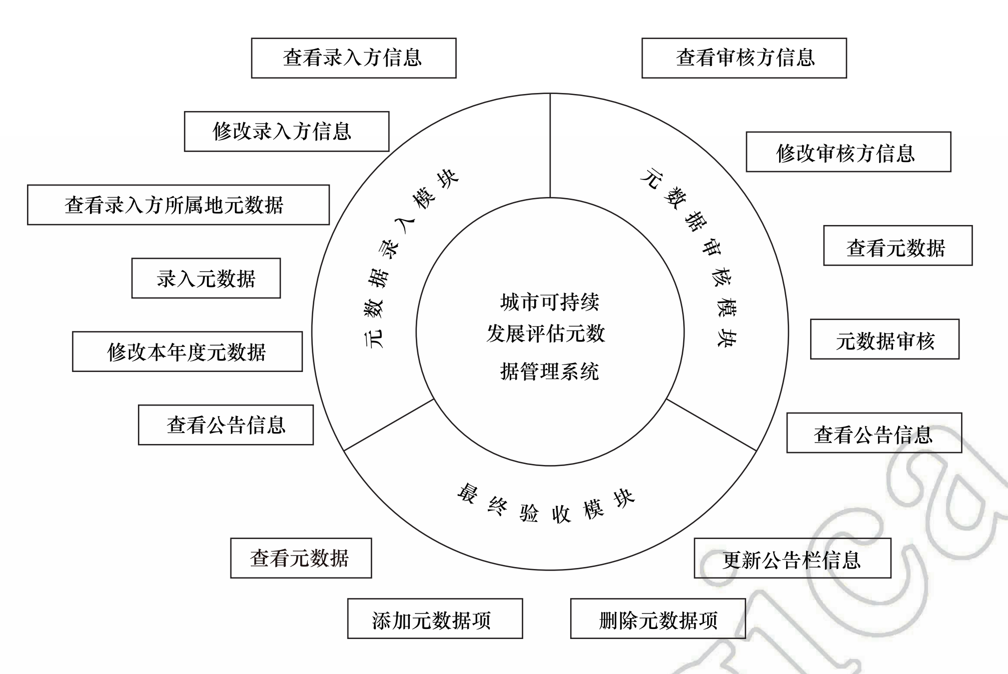


图 2 元数据管理系统组织结构

Fig.2 Organizational structure of metadata management system

同时,根据元数据管理系统的功能及相关填报要求,针对《城市可持续发展能力评估指标体系》的抽样指标进行填报,同时选取“生态系统服务价值总量”指标做填报示范(表 6)。

表 6 元数据管理系统填报示范

Table 6 Data entry demonstration of the metadata management system

元数据项名称 The metadata name	元数据信息 Metadata
数据项编号 Data item number	20
评估指标 Assessment index	生态系统服务价值总量
实验区 Experimental area	绍兴市
数据录入年份 Year of data entry	2015
数据发布日期 Data release date	2015-7-29
数据发布形式 Data release form	绍兴市科技信息研究院统计表
空间范围 Spatial range	市域
起始时间 Start time	2014-1-1
结束时间 End time	2014-12-31
统计频率 Statistical frequency	年
数据安全限制分级 Data security limit classification	公开
数据志说明 Data description	该指标统计结果基于《城市可持续发展能力评估指标体系》中的公式计算获取,参考数据中的生态系统类型的面积来自绍兴市国土局的遥感监测数据
在线资源链接地址 Online resource link address	无
数据统计单位名称 Data statistics institution	绍兴市科技信息研究院
数据统计单位联络人 Contacts of data statistics institution	徐先生
数据统计单位联系电话 Contact number of data statistics institution	—
数据统计单位地址 Address of data statistics institution	—
数据统计单位邮箱地址 E-mail address of data statistics institution	—

4 结论与展望

城市可持续发展能力是一个由经济、社会、资源和环境等要素构成的复杂系统,建立一个既能反映各要素

协调发展、又能衡量城市可持续发展状态的评价指标体系是开展城市可持续发展评估的前提。而数据作为填充评价指标体系这个骨架内容的血肉,其能否高质量的获取是衡量评估结果科学有效的关键。深入分析每一个评估数据的来源和生成方法及其所表征的空间范围、时间范围和城市发展属性,对于准确把握城市可持续发展能力具有重要意义。

从本文的研究结果看,元数据理论与技术在解决评估数据质量问题方面是一种行之有效的方法,并且研发使用元数据管理系统能更为便捷的管理和记录元数据,提高数据数据质量,从而保障城市可持续发展能力评估结果的科学有效。

展望未来,我国城市可持续发展能力评估工作应全面采用元数据管理技术,并促进元数据管理系统与可持续发展能力评估系统有效衔接,让每一个评估指标数据有源可循,有据可查,才能让评估结果具有更强的公信力。同时,应进一步完善相关法律法规对于城市可持续发展能力评估指标体系元数据信息的重视保障,提高各个委办局对于评估数据元数据重要性的认识。

参考文献(References):

[1] 刘耀彬. 武汉市城市可持续发展能力评估. 城市发展研究, 2002, 9(2): 67-71.

[2] 彭程, 陈志芬, 吴华瑞, 孙想, 姚娜. 基于 ESDA 的城市可持续发展能力时空分异格局研究. 中国人口·资源与环境, 2016, 26(2): 144-151.

[3] 孙亚南. 长江经济带核心城市可持续发展能力评价. 南京社会科学, 2016(8): 151-156.

[4] Wei Y, Huang C, Lam P T I, Sha Y, Feng Y, Rosen M A. Using Urban-Carrying Capacity as a Benchmark for Sustainable Urban Development: An Empirical Study of Beijing. SUSTAINABILITY-BASEL, 2015, 7(3): 3244-3268.

[5] 徐浩. 区域可持续发展能力评估--以湖北省十堰市为例[D]. 北京:中央民族大学,2011.

[6] 刘晓梅. 树立正确的统计数据质量概念刍议. 统计与信息论坛, 2003, 18(5): 17-18.

[7] 王磊. 统计数据质量内涵的深入分析. 中国电子商务, 2012(17): 138.

[8] 吴凡, 苗韧. 城市可持续发展能力评估体系构建研究. 生态经济, 2017, 33(3): 105-109.

[9] 武青海, 夏洪波. 大数据时代下政府部门信息资源共享策略研究. 统计与管理, 2015(11): 104-105.

[10] 龙军, 陈舜子, 刘高嵩. 科技评价系统中的元数据表示与管理. 计算机技术与发展, 2010, 20(10): 12-15.

[11] 汤敬谦, 杨鹤林. 热点、网络与态势:国外图书情报学领域元数据研究的知识图谱分析. 图书馆学研究, 2016(6): 15-26.

[12] Harvey M J, Mclean A, Rzepa H S. A metadata-driven approach to data repository design. J CHEMINFORMATICS, 2017, 9(1): 4.

[13] 孙晓, 刘旭升, 李锋, 陶宇. 中国不同规模城市可持续发展综合评价. 生态学报, 2016, 36(17): 5590-5600.

[14] 袁明. 统计数据质量问题研究. 科学之友, 2008(36): 92-93.

[15] 田仲, 李培军, 程芳. 通用数据质量评分系统的研究与设计. 标准科学, 2016(5): 94-99.

[16] 许从余. 土地利用数据质量控制与评价体系研究[D]. 杭州:浙江大学, 2011.

[17] Vidgen R. Constructing a web information system development methodology. Information Systems Journal, 2002, 12(3): 247-261.

[18] 张英朝, 邓苏, 张维明. 数据仓库元数据管理研究. 计算机工程, 2003, 29(1): 8-10.

[19] 诸葛剑平. 分级控制原理在行政管理体制改革中的运用. 行政论坛, 2002(1): 13-14.

[20] 李长征, 雷勇. 分级权限管理在数据库应用系统中的实现. 机械设计与制造, 2006(2): 171-172.